Exercise: System identification and bivariate ARMAX model

Instructions: This exercise should give hands-on experience of finding a good bivariate ARMAX model given data from an experiment – this will hopefully insights and inspiration for similar modelling tasks.

Data and a script for solving most of it is given in the accompanying `.zip` file: Open the `script.R` and for each question below the name of the chunk of code belonging to it is underlined.

Introduction

First a description of the experiment system, setup and data is given.

The system consists of a pot and a drinking glass, both filled with water. The drinking glass is placed in the pot, such that the water of the pot covers up to half the height of the glass. In both the pot and the glass is placed a 300 W water heater.

In Figure 1 two pictures of the system are shown. As it can be seen a small "model-ship outboard motor" was used to mix up the water, such that the water is mixed and the temperature in the each compartment can be assumed homogenous with the applied sampling frequency. In Figure 2 a system diagram depicts the setup.

The following variables was measured:

- $T_{\mathrm{i},t}$ (`Tinner`) (°C): The temperature of the water in the glass
- $T_{\mathrm{o},t}$ (`Touter`) (°C): The temperature of the water in the pot
- $T_{\mathrm{a},t}$ (`Ta`) (°C): The temperature of the ambient air surrounding the system
- $P_{\mathrm{i},t}$ (`Pinner`) (W): The Power input of the heater in the glass
- $P_{\mathrm{o},t}$ (`Pouter`) (W): The power input of the heater in the pot

The variables are sampled with a sampling period of 10 second. In the data the column `t` is Unix time UTC.

Data from a single experiment is used. In the experiment both heaters were switched on and off following to a designed test sequence, such that the on-off periods were independent.

Figure 1: In the left image the setup is seen from the side and in the right image from above (the right image don't have the mixing moter, but it was in the experiment used for the assignment).

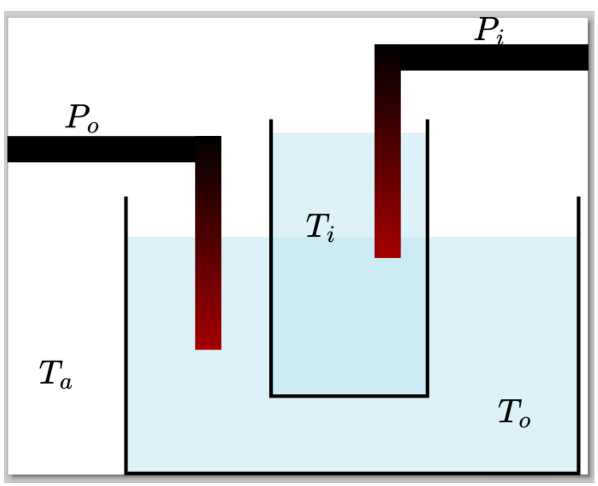

Figure 2: Diagram of the experiment. Each compartment can be assumed homogeneous meaning that the temperature is the same in all of the material (water in this case) in the compartment.

1 Explorative data analysis

1.1. ReadAndPlot
Make plots of the data in `data/experiment1.csv`.

1.2. From the plots describe the experiment. What are the causalities, i.e. which variables are "dependent" and which are "independent"? Is the system a linear time-invariant (LTI) system or what could cause it not to be?

1.3. CCFs
Try to investigate the relations between the variables using the cross-correlation function between the variables. Do the CCFs provide more insights than the time series plots?

2 ARX model of the glass water temperature

First, we want to build a model with the drinking glass water temperature (`Tinner`) as output and ambient temperature as input.

Start with an ARX model only with the ambient temperature

$$\phi(B)T_{\mathrm{i},t} = \omega_{\mathrm{a}}(B)T_{\mathrm{a},t} + \varepsilon_t \tag{1}$$

where ε_t is assumed i.i.d. Keep the same order of AR and the all the coefficient polynomials.

For example the ARX model above of order 1 written out as a regression model is

$$T_{\mathrm{i},t} = -\phi_1 T_{\mathrm{i},t-1} + \omega_{\mathrm{a},1}T_{\mathrm{a},t-1} + \varepsilon_t \tag{2}$$

and of order 2

$$T_{\mathrm{i},t} = -\phi_1 T_{\mathrm{i},t-1} - \phi_2 T_{\mathrm{i},t-2} + \omega_{\mathrm{a},1}T_{\mathrm{a},t-1} + \omega_{\mathrm{a},2}T_{\mathrm{a},t-2} + \varepsilon_t \tag{3}$$

and so on for higher orders.

2.1. <u>ARX1</u>
Estimate the ARX order 1 model in Equation (2) as a linear regression model and estimate the parameters with the LS method. Make a validation, and think about if you want to add more inputs: if yes, add iteratively until you have a the 1. order model.

As you see in the R script you can make the lags with the provided `lagdf()` function and use the `lm()` function and use the provided `ARX()` function to easily generate a formula for higher order models.

2.2. <u>ARX-select</u> and <u>ARX-select-AICBIC</u>
Identify a suitable ARX model for $T_{\mathrm{i},t}$. Add relevant inputs and increase the order (use same for all inputs and AR, i.e. always the same number of lags (from 1 to "order") for all variables).

Argue for the selected inputs and order of the suitable model (i.e. neither under- nor over-fitted).

Tip: In general, as experience grows, one tends to become more "conservative" (as with other things in life ;) and choose simpler models. One argument, that usually many other effects come into play, e.g. some phenomena not in the data, etc.

Tip: In general, it's usually a good idea to first find out which inputs to use and then start increase the order of the model, but there is no final correct way, but try to be structured following some kind of model selection procedure.

2.3. <u>ARX-multistep-prediction</u>
Run a multi-step prediction

3 ARMAX model of the glass water temperature

Install the marima2 package from a downloaded file, note it is a slightly modified version of the original package:

```r
# Install a slightly
download.file("https://02417.compute.dtu.dk/material/marima2_0.1.tar.gz", "marima2_0.1.tar.gz")
install.packages("marima2_0.1.tar.gz", repos=NULL)
library(marima2)
```

Note, the way the model is specified (in the modified package) is e.g.:

```r
fit <- marima("Tinner ~ AR(1) + Pinner(1) + Touter(1) + Ta(1)", data=X)
```

The syntax for the model formula is e.g.:

```r
fit <- marima("Y ~ AR(1:2) + x1(1:2) + x2(1:2) + MA(1:2)", data=X)
```

would give a 2. order ARMAX.

For each input the lags can be set independent, e.g.:

```r
fit <- marima("Y ~ AR(2:3) + x(1) + MA(4)", data=X)
```

would give

$$Y_t = \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \omega_1 x_{t-1} + \varepsilon_t + \theta_4 \varepsilon_{t-4}$$

3.1. ARMAX-select
Identify an ARMAX model. Increase the order (same order for AR and all inputs), use same techniques as for the ARX from last section for model selection and validation.

However, note that AIC and BIC are not calculated on exactly the same observations, because of NAs due to lags. So pool all the info from significance, validation of residuals, multi-step prediction, to argue for the choice.

3.2. ARMAX-validate
Discuss the results, did you get the same results or is there some difference between the results for ARX and ARMAX?

3.3. ARMAX-multistep
Make a multi-step prediction with the selected model. Can it predict the temperature through out the exeriment?

4 ARMAX bivariate model

4.1. ARMAXbivariate-select
Use marima to estimate an ARMAX bivariate model of the coupled system. Use the same model inputs and order as identified in the individual models.

4.2. Identify the order. Increase or decrease the order for all inputs together, to the extend possible use the same techniques as before.

4.3. ARMAXbivariate-multistep-prediction
Make a multi-step prediction with the selected model. Can it predict both temperatures through out the exeriment?

4.4. ARMAXbivariate-stepresponse
Simulation of step response from one input to both temperatures simulatiously. E.g. step 100 W on $P_{\mathrm{i},t}$ and see the effect on both temperatures, play around, and present three interesting simulations.

4.5. Discuss the pros and cons of the coupled model over the two independent models.

4.6. ARMAXbivariate-multistep-with-new-power-sequence
We can simply use the model to predict with any inputs, this could be used for Model Predictive Control (look it up), which is a typical use of such models!

4.7. ARMAXbivariate-steady-state-response
It's possible to calculate the steady-state response for an input to both the outputs.